

# Deep learning for ICD coding: Looking for medical concepts in clinical documents in english and in French

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

---

## Abstract

© Springer Nature Switzerland AG 2018. Medical Concept Coding (MCD) is a crucial task in biomedical information extraction. Recent advances in neural network modeling have demonstrated its usefulness in the task of natural language processing. Modern framework of sequence-to-sequence learning that was initially used for recurrent neural networks has been shown to provide powerful solution to tasks such as Named Entity Recognition or Medical Concept Coding. We have addressed the identification of clinical concepts within the International Classification of Diseases version 10 (ICD-10) in two benchmark data sets of death certificates provided for the task 1 in the CLEF eHealth shared task 2017. A proposed architecture combines ideas from recurrent neural networks and traditional text retrieval term weighting schemes. We found that our models reach accuracy of 75% and 86% as evaluated by the F-measure on the CépiDc corpus of French texts and on the CDC corpus of English texts, respectfully. The proposed models can be employed for coding electronic medical records with ICD codes including diagnosis and procedure codes.

[http://dx.doi.org/10.1007/978-3-319-98932-7\\_19](http://dx.doi.org/10.1007/978-3-319-98932-7_19)

---

## Keywords

CDC corpus, CLEF eHealth, Computer assisted coding, CépiDc corpus, Death certificates, Deep learning, Encoder-decoder model, Healthcare, ICD codes, ICD coding, Machine learning, Medical concept coding, Medical record coding, Recurrent neural network

## References

- [1] Pradhan, S., Elhadad, N., Chapman, W.W., Manandhar, S., Savova, G.: SemEval-2014 task 7: analysis of clinical text. In: SemEval@ COLING, pp. 54-62 (2014)
- [2] Dougherty, M., Seabold, S., White, S.E.: Study reveals hard facts on CAC. J. AHIMA 84(7), 54-56 (2013)
- [3] Stanfill, M.H., Williams, M., Fenton, S.H., Jenders, R.A., Hersh, W.R.: A systematic literature review of automated clinical coding and classification systems. J. Am. Med. Inform. Assoc. 17(6), 646-651 (2010)
- [4] Miftahutdinov, Z., Tutubalina, E.: KFU at CLEF ehealth 2017 task 1: ICD-10 coding of English death certificates with recurrent neural networks. In: CEUR Workshop Proceedings, vol. 1866 (2017)
- [5] Karimi, S., Dai, X., Hassanzadeh, H., Nguyen, A.: Automatic diagnosis coding of radiology reports: a comparison of deep learning and conventional classification methods. In: BioNLP 2017, pp. 328-332 (2017)
- [6] Duarte, F., Martins, B., Pinto, C.S., Silva, M.J.: Deep neural models for ICD-10 coding of death certificates and autopsy reports in free-text. J. Biomed. Inform. 80, 64-77 (2018)
- [7] Zhang, Y., et al.: Uth CCB: a report for SemEval 2014-task 7 analysis of clinical text. In: SemEval 2014, p. 802 (2014)

- [8] Ghiasvand, O., Kate, R.J.: UWM: disorder mention extraction from clinical text using CRFS and normalization using learned edit distance patterns. In: SemEval@ COLING, pp. 828-832 (2014)
- [9] Van Mulligen, E., Afzal, Z., Akhondi, S.A., Vo, D., Kors, J.A.: Erasmus MC at CLEF eHealth 2016: concept recognition and coding in French texts. In: CLEF (2016)
- [10] Cabot, C., Soualmia, L.F., Dahamna, B., Darmoni, S.J.: SIBM at CLEF eHealth evaluation lab 2016: extracting concepts in French medical texts with ECMT and CIMIND. In: CLEF (2016)
- [11] Mottin, L., Gobeill, J., Mottaz, A., Pasche, E., Gaudinat, A., Ruch, P.: BiTeM at CLEF eHealth evaluation lab 2016 task 2: multilingual information extraction. In: CEUR Workshop Proceedings, vol. 1609, pp. 94-102 (2016)
- [12] Dermouche, M., Looten, V., Flicoteaux, R., Chevret, S., Velcin, J., Taright, N.: ECSTRA-INSERM@ CLEF eHealth2016-task 2: ICD10 code extraction from death certificates. In: CLEF (2016)
- [13] Zweigenbaum, P., Lavergne, T.: LIMSI ICD10 coding experiments on CépiDC death certificate statements. In: CLEF (2016)
- [14] Leaman, R., Khare, R., Lu, Z.: NCBI at 2013 shARe/CLEF ehealth shared task: disorder normalization in clinical notes with DNorm. *Radiology* 42(21.1), 1-941 (2011)
- [15] Suominen, H., et al.: Overview of the ShARe/CLEF eHealth evaluation lab 2013. In: Forner, P., Müller, H., Paredes, R., Rosso, P., Stein, B. (eds.) CLEF 2013. LNCS, vol. 8138, pp. 212-231. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-40802-1\\_24](https://doi.org/10.1007/978-3-642-40802-1_24)
- [16] Leaman, R., Islamaj Doğan, R., Lu, Z.: DNorm: disease name normalization with pairwise learning to rank. *Bioinformatics* 29(22), 2909-2917 (2013)
- [17] Névél, A., et al.: CLEF ehealth 2017 multilingual information extraction task overview: ICD10 coding of death certificates in English and French. In: CLEF 2017 Evaluation Labs and Workshop: Online Working Notes, CEUR-WS (2017)
- [18] Névél, A., et al.: Clinical information extraction at the CLEF eHealth evaluation lab 2016. In: Proceedings of CLEF 2016 Evaluation Labs and Workshop: Online Working Notes, CEUR-WS, September 2016 (2016)
- [19] Zweigenbaum, P., Lavergne, T.: Hybrid methods for ICD-10 coding of death certificates. In: EMNLP 2016, p. 96 (2016)
- [20] Cabot, C., Soualmia, L.F., Darmoni, S.J.: SIBM at CLEF ehealth evaluation lab 2017: multilingual information extraction with CIM-IND. In: CLEF (2017)
- [21] Tutubalina, E., Miftahutdinov, Z., Nikolenko, S., Malykh, V.: Medical concept normalization in social media posts with recurrent neural networks. *J. Biomed. Inform.* 84, 93-102 (2018)
- [22] Rios, A., Kavuluru, R.: EMR coding with semi-parametric multi-head matching networks. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), vol. 1, pp. 2081-2091 (2018)
- [23] Schuemie, M.J., Kors, J.A., Mons, B.: Word sense disambiguation in the biomedical domain: an overview. *J. Comput. Biol.* 12(5), 554-565 (2005)
- [24] Névél, A., et al.: CLEF eHealth 2017 Multilingual information extraction task overview: ICD10 coding of death certificates in English and French. In: Working Notes of Conference and Labs of the Evaluation (CLEF) Forum, CEUR Workshop Proceedings (2017)
- [25] Elman, J.L.: Finding structure in time. *Cogn. Sci.* 14(2), 179-211 (1990)
- [26] Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* 5(2), 157-166 (1994)
- [27] Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J.: LSTM: a search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* 28(10), 2222-2232 (2016)
- [28] Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078* (2014)
- [29] Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. *IEEE Trans. Sig. Proc.* 45(11), 2673-2681 (1997)
- [30] Graves, A., Fernández, S., Schmidhuber, J.: Bidirectional LSTM networks for improved phoneme classification and recognition. In: Duch, W., Kacprzyk, J., Oja, E., Zadrozny, S. (eds.) ICANN 2005. LNCS, vol. 3697, pp. 799-804. Springer, Heidelberg (2005). [https://doi.org/10.1007/11550907\\_126](https://doi.org/10.1007/11550907_126)
- [31] Graves, A., Schmidhuber, J.: Framewise phoneme classification with bidirectional LSTM networks. In: Proceedings of IEEE International Joint Conference on Neural Networks, IJCNN 2005, vol. 4, pp. 2047-2052. IEEE (2005)
- [32] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems, pp. 3111-3119 (2013)
- [33] Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. *Inf. Proc. Manag.* 24(5), 513-523 (1988)

- [34] Miftahutdinov, Z., Tutubalina, E., Tropsha, A.: Identifying disease-related expressions in reviews using conditional random fields. In: Proceedings of International Conference on Computational Linguistics and Intellectual Technologies Dialog, vol. 1, pp. 155–167 (2017)
- [35] Bojanowski, P., Grave, E., Joulin, A., Mikolov, T.: Enriching word vectors with subword information. Trans. Assoc. Comput. Linguist. 5, 135–146 (2017)
- [36] Moen, S., Ananiadou, T.S.S.: Distributional semantics resources for biomedical text processing (2013)
- [37] Chollet, F., et al.: Keras (2015). <https://github.com/fchollet/keras>
- [38] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15(1), 1929–1958 (2014)
- [39] Kinga, D., Adam, J.B.: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR) (2015)